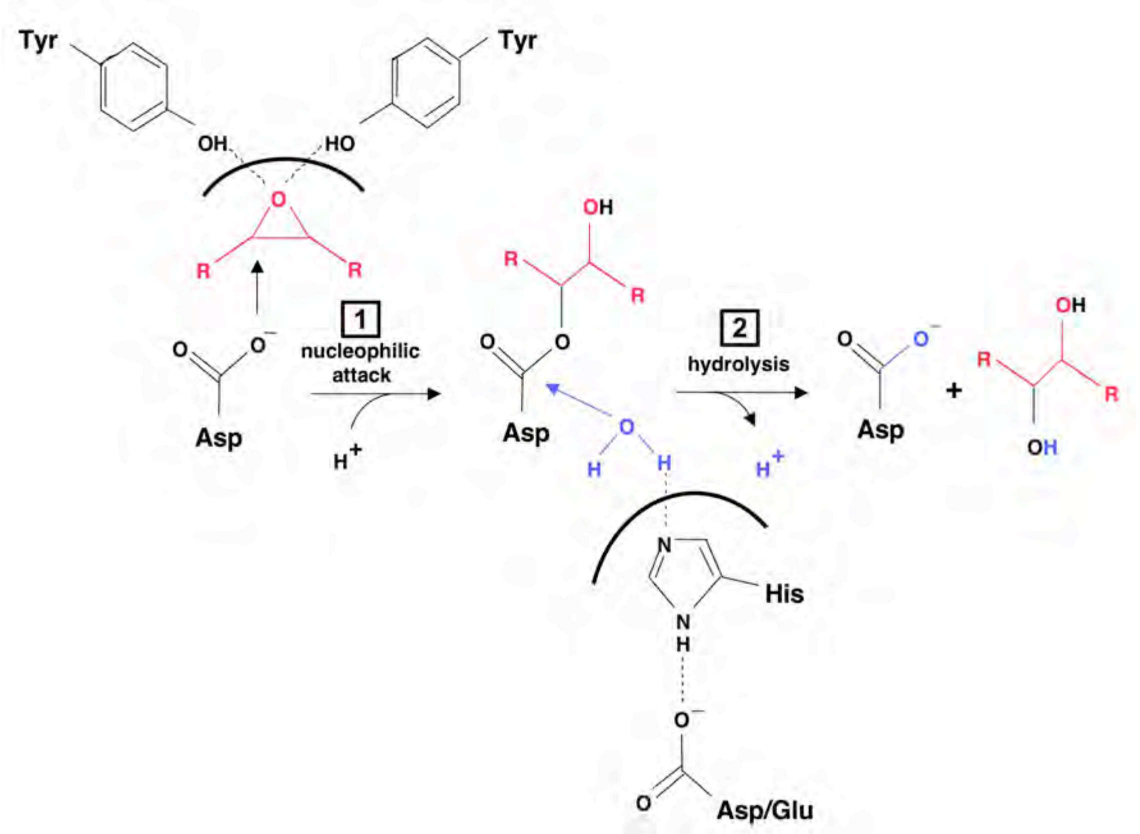
Supporting Information 1

Enzymatic mechanism of α/β hydrolase fold epoxide hydrolases



The active site of α/β hydrolase fold epoxide hydrolases is situated at the interface of the two essential domains of these proteins, i) the so-called α/β hydrolase fold domain and ii) the lid domain, which is sitting on top of the former. Essentially five amino acid residues are involved in the catalysis. Two tyrosine residues reaching from the lid domain into the substrate binding pocket form hydrogen bonds to the epoxide oxygen of the substrate molecule that enters the pocket. This keeps the substrate in the proper spatial position and also activates it for the first enzymatic step, the nucleophilic attack. The catalytic nucleophile, invariably an aspartic acid residue, forms an ester bond to one of the epoxide ring carbon atoms, while the oxygen-carbon bond in the ring opens under formation of a hydroxy group adjacent to the ester bond. The required proton is provided by one of the tyrosine residues. The second catalytic step, the hydrolysis, is catalyzed by a charge relay system composed of a histidine residue and an acidic residue, again an aspartic acid residue in many EHs. The charge relay system activates a water molecule by proton abstraction that, in turn, hydrolyzes the ester bond between enzyme and substrate through attack as a hydroxyl anion at the carboxylate carbon of the catalytic nucleophile. As a result, the product of the reaction, a so-called vicinal diol, is released and the enzyme is regenerated. For historical reasons, catalytic nucleophile and charge relay system are collectively termed the catalytic triad of the enzyme because the essential function of the tyrosine residues has become obvious only much later.

Supporting Information 2

Sequence analysis of the novel putative epoxide hydrolases

The first sequence comparison of epoxide hydrolases (Arand *et al.* (1994) Sequence Similarity of Mammalian Epoxide Hydrolases to the Bacterial Haloalkane Dehalogenase and Other Related Proteins - Implication for the Potential Catalytic Mechanism of Enzymatic Epoxide Hydrolysis. *FEBS Lett* 338: 251-256) revealed two neighbouring hot spots of sequence conservation, 6 and 5 amino acids in length. Further analysis (Arand *et al.* (1996) in *Control Mechanisms of Carcinogenesis*, eds. Hengstler and Oesch, pp. 116-134) allowed to combine and extend these (underlined) to a single, continuous 16 amino acid motif, <u>RVIAPD</u>LR<u>GYGDS</u>DKP, with an average degree of conservation of 60% per residue. The candidates obtained with the BLASTP search of the human "build protein" database using the above peptide sequence as the bait were further screened for the presence of the following motifs:

1) a H-G-X-P tetrapeptide about 15-30 aa N-terminal to the location that matched the bait sequence; this is a conserved sequence, forming part of the oxyanion hole, a structural feature that stabilizes the tetrahedral intermediate formed during the second, hydrolytic step of the catalysis; in epoxide hydrolases, X is usually an aromatic residue, often a tryptophane, while dehalogenases, the closest relatives of EHs, usually have a glutamic acid residue in this position

2) a sm-X-D-hy-sm-sm hexapeptide 28-32 aa C-terminal to the location that matched the bait sequence; this is the motif containing the catalytic nucleophile, invariably a D in epoxide hydrolases and dehalogenases, and an S in most other α/β hydrolase fold proteins; sm needs to be a small aa residue, due to sterical constraints in the structure of the catalytic elbow, X is not clearly defined but often a histidine in epoxide hydrolases and hy is usually a hydrophobic amino acid, often a tryptophane in epoxide hydrolases

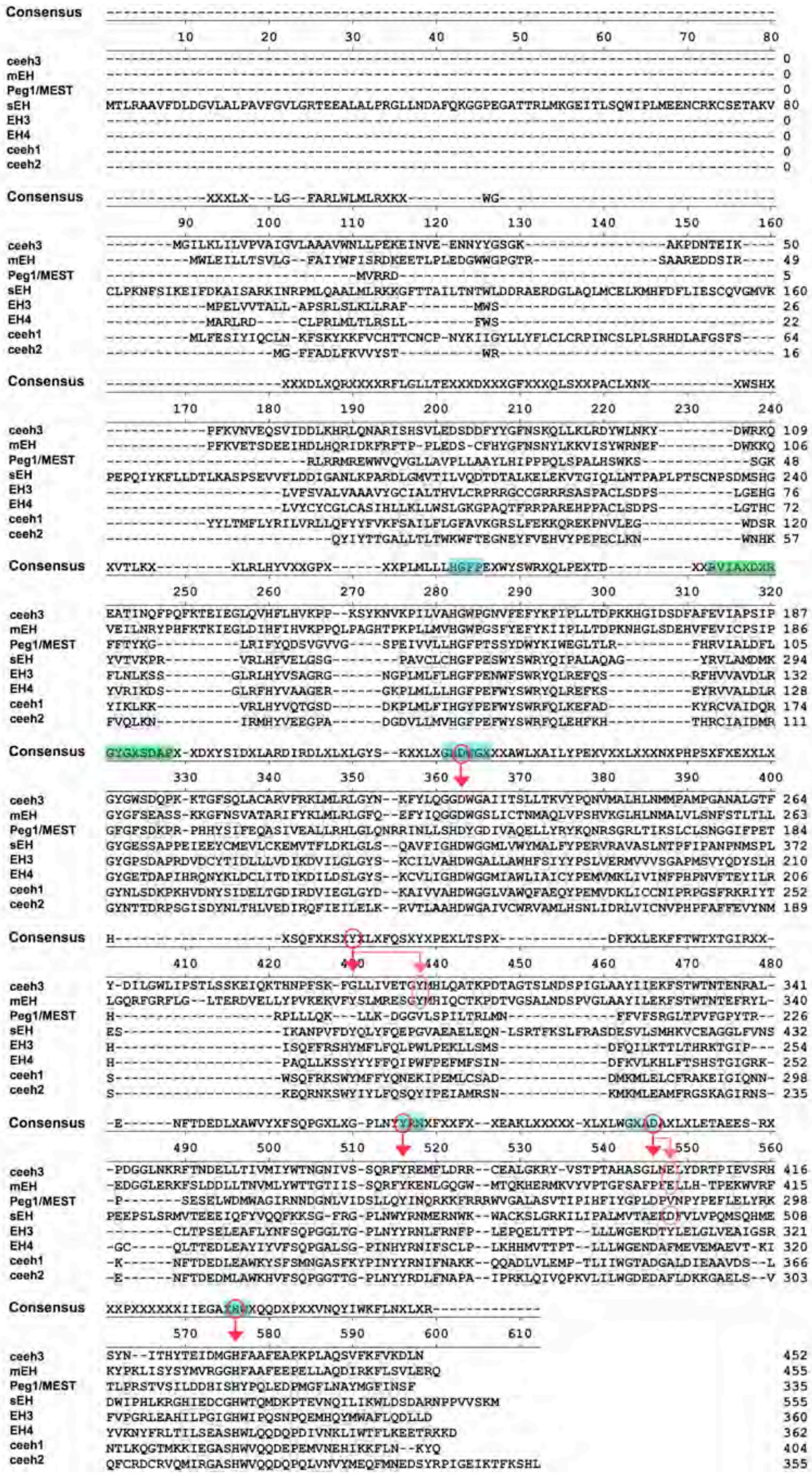
3) two tyrosines in the region of the potential lid domain, in a distance of 50-85 aa from each other; unfortunately, the sequence context of the first tyrosine is not at all conserved; the second one is located in a tripeptide with the consensus sequence Y-R-N, where R may be replaced by K, and N may be replaced by D or E, but other substitutions occur as well; these two tyrosines are involved in the initial substrate recognition and activation of the oxiran ring and are absent in dehalogenases and esterases

4) a poorly conserved tetrapeptide G-E-L-D, roughly 30 aa C-terminal to the second tyrosine; the D (or E in case of mEH-related enzymes) is the acidic residue of the charge relay system that catalyzes the hydrolysis of the covalent intermediate, the G is the second best conserved residue within this motif; this motif is not conserved in mEH-related EHs, and D is replaced by E; in some microbial EHs, the localization of the acidic residue in the amino acid sequence is apparently not conserved

5) a moderately conserved heptapetide G-H-W-T-Q-I-E, ≥ 20 aa upstream of the C-terminus; H is the calatytic histidine, affording the water activation in the charge relay system

The position of these motifs is indicated in the sequence comparison below (Fig. SI2), that also forms the basis for the phylogenetic tree in Fig. 1 of the manuscript, as well as for the selection of candidate catalytic residues that were mutagenized.

Fig. SI2



Supporting Information 3

Rational of the nomenclature for the new epoxide hydrolases

The family of α/β hydrolase fold epoxide hydrolases in mammals appears to be small, according to our present analysis. Two hitherto known members, mEH and sEH, and two to three new members (EH3 and EH4 and, possibly peg1/MEST, then to be renamed to EH5) result in a low complexity. This made as deviate from two possible nomenclature alternatives and keep the designation most simple by using consecutive numbers.

One of the obvious alternatives would have been to follow the more complex nomenclature system of, e.g. the CYP enzyme superfamily (Nelson (2006) Cytochrome P450 nomenclature, 2004. *Methods Mol Biol* 320: 1-10) that indicates the degree of similarity by sorting the members into subfamilies within families. However, with presently not more than 5 enzymes in total and only one family having more than 1 member, this seemed unnecessary.

We also did not adapt the accurate, complex scheme recently established for epoxide hydrolases from all phylae (van Loo et al. (2006) Diversity and biocatalytic potential of epoxide hydrolases identified by genome analysis. *Applied and Environmental Microbiology* 72: 2905-2917.), according to which EH3 and EH4 belong to group 2, mEH belongs to group 5 and sEH belongs to group 8 of the α/β hydrolase fold epoxide hydrolases, because this would lead to significant confusion within the large community of researchers with an exclusive interest in human/mammalian EHs.

Supporting Information 4

Primer and PCR conditions for the generation of EH3 mutants

pFastBac EH3 served as the template for the mutagenesis PCR. Reaction set up was composed as follows: 10-30 ng plasmid, forward and reverse primer (see Tab. SI4), 200 ng each, dNTPs, 250 µM each, and 1 µl Pfu Ultra™ HF DNA Polymerase in a final volume of 50 µl Pfu Ultra HF reaction buffer (Stratagene, La Jolla, CA).

The PCR reaction was started with the initial denaturation at 95 °C for 3 min, followed by 30 cycles of: (1) denaturation for 40 s at 95 °C, (2) annealing for 30 seconds (for individual temperatures used for each primer pair see Tab. SI4), and (3) extension for 6 min at 72 °C. The reaction was completed with a final extension at 72 °C for 10 min. Subsequently, the PCR products were treated with *Dpn* I (one hour at 37 °C) to digest the methylated wild-type parental plasmid. Finally, E.coli DH10B were transformed with 1 µl of each reaction and plasmids isolated from individual colonies obtained after plating were analyzed for the presence of the desired mutations and the absence of any other sequence alterations. The final constructs obtained this way were further processed to obtain recombinant baculoviruses as described in Materials and Methods.

Tab. SI4 Primer sequences and annealing temperatures

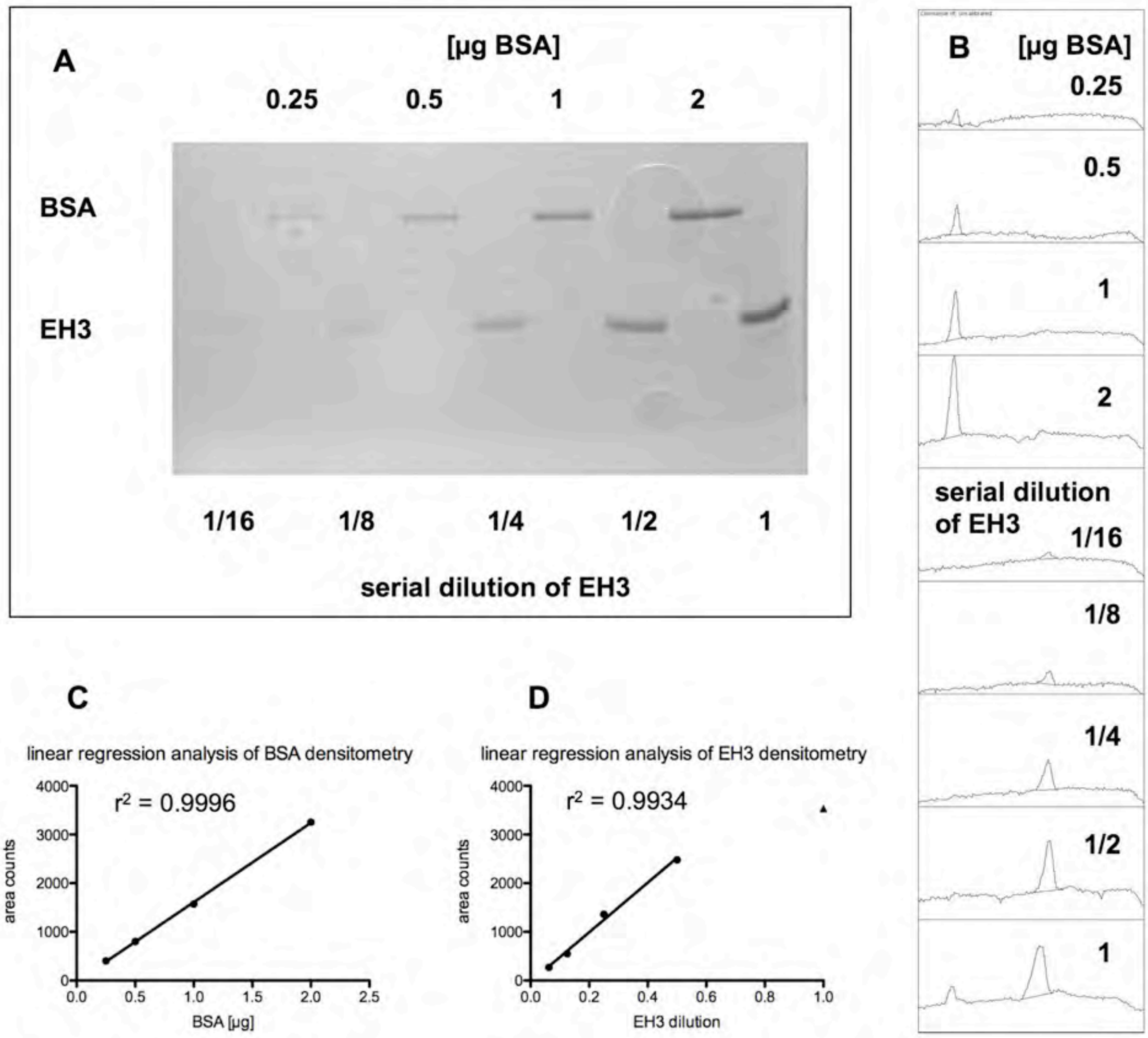
Amino acid substitution	PCR-primers	Annealing temperature [°C]
D173A	forward: GTGGCCCATGCCTGGGGTGCCCT reverse: AGGAGGGCACCCCAGGCATGGG	66
D173N	forward: GTGGCCCATAACTGGGGTGCCCT reverse: AGGAGGGCACCCCAGTTATGGG	62
Y220F	forward: CCGTTCCCACTTCATGTTCCTG reverse: GCTGGAACAGGAACATGAAGTGGG	58
Y280F	forward: CCTCAACTTCTACCGAAACC reverse: CTGAAGAGGTTTCGGTAGAAGTTG	54
Y281F	forward: CCTCAACTACTTCCGAAACC reverse: CTGAAGAGGTTTCGGAAGTAGTTG	54
D307A	forward: GGAGAAGGCCACTTACTTGGAGC reverse: GCTCCAAGTAAGTGGCCTTCTCC	60
D307N	forward: GCTGTGGGGGGAGAAGAACACT reverse: GCTCCAAGTAAGTGTTCTTCTCC	56
H337A	forward: AGGCATAGGGGCTTGGATCCCAC reverse: TGCTCTGTGGGATCCAAGCCCCT	62
H337Q	forward: GCATAGGGCAATGGATCCCACAG reverse: GCTCTGTGGGATCCATTGCCC	60

Supporting Information 5

Immunoquantification of EH3 in insect cell lysates

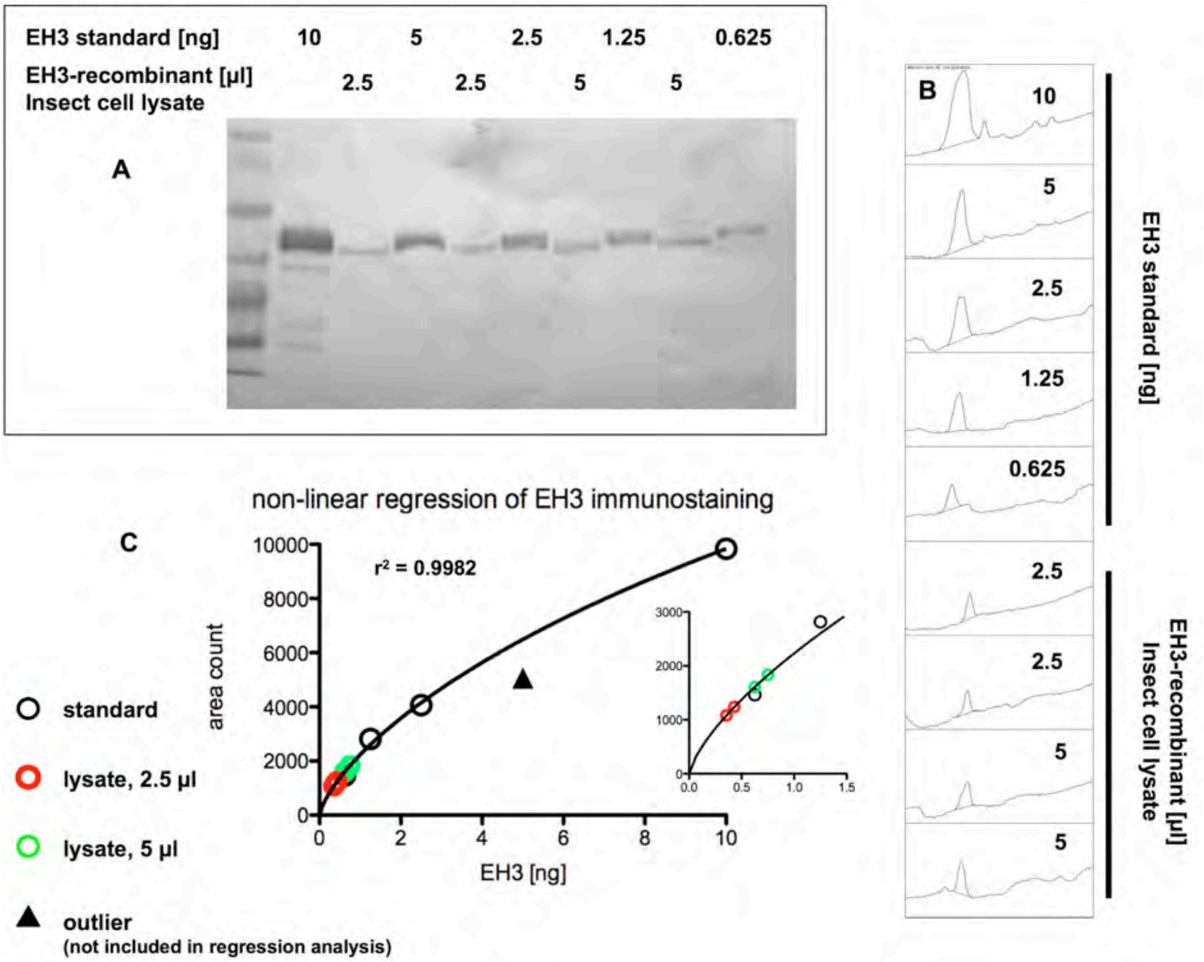
For the calculation of the kinetic properties of EH3, the enzyme was immunoquantified in insect cell lysates. In a first step, recombinant EH3 obtained as inclusion bodies via overexpression in E.coli was washed in Tris-buffered saline (Tris-HCl, 20 mM, NaCl, 100 mM, pH 8.0), inclusion bodies were recovered by brief centrifugation in a benchtop centrifuge and dissolved in high SDS sample buffer (sodium dodecyl sulfate, 5 %, glycerol, 8 %, 2-mercaptoethanol, 3.3 %) to yield a final concentration in the range of 0 – 0.3 µg/µl. Ten microliters of 1:1 serial dilutions of the dissolved EH3 protein were analyzed via SDS-PAGE and subsequent Coomassie blue staining, and compared to samples with known, increasing concentrations (0.25 – 2 µg per lane) of a bovine serum albumine that were run along. Gels were photographed with the digital camera of a gel documentation system (GeneFlash®, Syngene, Cambridge, UK) and the band intensities were quantified with the public domain software package ImageJ (NIH, USA), using the routines of the *Analyze/Gels* sub menue (Fig. SI5-1). This way, an immunological standard of known EH3 concentration was established.

Fig. SI5-1



Using this standard, the recombinant EH3 present in insect cell lysates was quantified by comparative immunoblotting, again using the densitometric features of ImageJ (Fig. SI5-2).

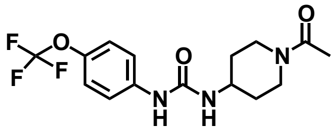
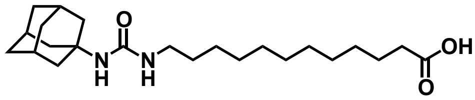
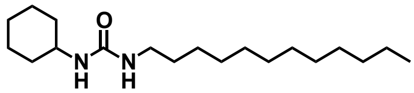
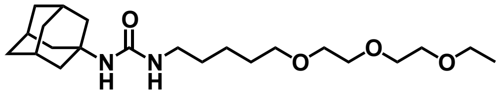
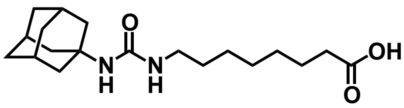
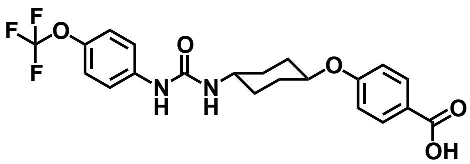
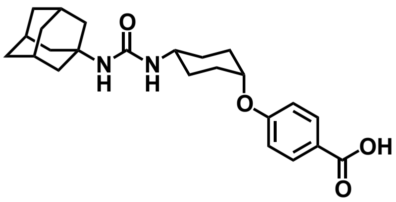
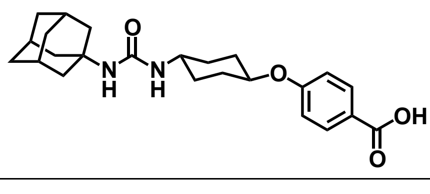
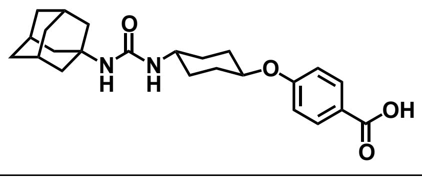
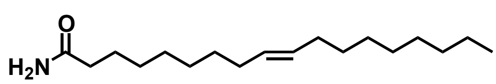
Fig. SI5-2



The Coomassie-stained polyacrylamid gel (SI5-1) and the immunoblot (SI5-2) were obtained as digital image (A) and densitometry was performed simulatenously on all lanes of the same image (B). Standard curves were obtained by linear (SI5-1) or non-linear (SI5-2) regression (C) using *Prism 5* and compared to either dilution series also analyzed by linear regression in the case of Coomassie-stained standard(SI5-1 D) or by direct reading from the standard curve in the case of immunoblot analysis (SI5-2 C). In the present example, the resulting concentration obtained for EH3 in the analyzed insect cell lysate was 0.15 ± 0.02 µg/ml (mean ±SD).

Supporting Information 6

Structure and inhibitory potency of EH inhibitors employed in the study

IUPAC name (common acronym)	structure	Inhibitory potency
1-(1-acetylpiperidin-4-yl)-3-(4-(trifluoromethoxy)phenyl) urea (TPAU)		+++
12-(3-((3*S*,5*S*,7*S*)-adamantan-1-yl)ureido) dodecanoic acid (AUDA)		+++
1-cyclohexyl-3-dodecylurea (CDU)		+++
1-((3*S*,5*S*,7*S*)-adamantan-1-yl)-3-(5-(2-(2-ethoxyethoxy) ethoxy)pentyl)urea (AEPU)		++
8-(3-((3*S*,5*S*,7*S*)-adamantan-1-yl)ureido)octanoic acid (AUOA)		+
4-(((1*R*,4*R*)-4-(3-(4-(trifluoromethoxy)phenyl)ureido)cyclohexyl)oxy)benzoic acid (t-TAUCB)		-
4-(((1*S*,4*S*)-4-(3-((3*S*,5*S*,7*S*)-adamantan-1-yl)ureido) cyclohexyl)oxy)benzoic acid (c-AUCB)		-
4-(((1*R*,4*R*)-4-(3-((3*S*,5*S*,7*S*)-adamantan-1-yl)ureido)cyclohexyl)oxy)benzoic acid (t-AUCB)		-
1-((3*S*,5*S*,7*S*)-adamantan-1-yl)-3-cyclohexylurea (ACU)		-
(*E*)-octadec-9-enamide (Elaidamide)		-